

Uso de inteligencia artificial generativa para la resolución de tareas complejas

Eva Millán
E.T.S. de Ingeniería Informática
Universidad de Málaga
España
emillan@uma.es

Juan José Escribano Otero
Escuela Superior Politécnica
Universidad Francisco de Vitoria
España
juanjose.escribano@ufv.es

María José García García
Estudios de Informática y Multimedia
Universitat Oberta de Catalunya
España
mjose.garcia.garcia@gmail.com

Resumen— Este artículo describe una ficha de actividad para utilizar inteligencia artificial generativa en el proceso de enseñanza y aprendizaje. La ficha, adaptable a diversas disciplinas, proporciona un marco estructurado para que el profesorado diseñe actividades y guíe al estudiantado en un uso efectivo de herramientas de inteligencia artificial generativa. A modo ilustrativo, se describe la aplicación específica de dicha ficha para diseñar una actividad de resolución de tareas complejas, que posteriormente se concreta para diseñar una actividad de aprendizaje del modelado con redes bayesianas. Para evaluar su efectividad, se realizó un análisis comparativo de las puntuaciones obtenidas en esta pregunta del examen en tres cohortes, observándose una mejora significativa en las calificaciones. Los resultados demuestran el impacto positivo del uso de estas herramientas en el aprendizaje y su potencial para fomentar un aprendizaje personalizado en contextos educativos avanzados.

Abstract— *This article describes an activity sheet for using generative artificial intelligence in the learning process. The activity sheet, adaptable to various disciplines, provides a structured framework for teachers to design activities and guide students in the effective use of generative Artificial Intelligence tools. As an illustration, the specific application of this sheet is described to design an activity for solving complex tasks, which is later concretized to design a learning activity for modeling with Bayesian networks. To evaluate its effectiveness, a comparative analysis of the grades obtained by the students across three cohorts was conducted, showing a significant improvement in grades. The results demonstrate the positive impact of using these tools on learning and their potential to foster personalized learning in advanced educational contexts.*

Palabras clave—Inteligencia artificial generativa, modelado con redes bayesianas, experiencia en el aula

I. INTRODUCCIÓN

En los últimos años, la inteligencia artificial (IA) generativa ha emergido como una herramienta poderosa en diversos campos, incluyendo la educación. Modelos como ChatGPT y otros han demostrado su capacidad para generar texto, resolver problemas y asistir en tareas complejas, ofreciendo un gran potencial para transformar la manera en que el estudiantado enfrenta desafíos académicos [1]. Sin embargo, el uso efectivo de estas herramientas en el aula requiere estrategias pedagógicas claras y estructuradas que les permitan aprovechar al máximo sus capacidades mientras desarrollan habilidades críticas y analíticas [2]. El uso de la IA generativa en la educación universitaria representa un gran reto para los equipos docentes, ya

que, siendo una oportunidad única para potenciar el aprendizaje del estudiantado, conlleva una serie de riesgos importantes. Un uso adecuado puede tener ventajas tales como proporcionar asistencia personalizada y adaptable, ya que permiten abordar de manera más eficiente tareas que tradicionalmente han sido complejas y demandantes [3]. Además, promueven la exploración de nuevas perspectivas y enfoques al resolver problemas, enriqueciendo la experiencia de aprendizaje.

Sin embargo, en un contexto en el que las habilidades tecnológicas y el pensamiento crítico son esenciales, un uso inadecuado puede hacer que el estudiantado las use como un “atajo” hacia la entrega de actividades evaluables, pero no necesariamente hacia el aprendizaje significativo, profundo y duradero. En [4] se discuten y comparten experiencias sobre la aplicación de la IA en el ámbito educativo, indicando algunos de sus riesgos y muchas de sus oportunidades. También en [5], dentro de ediciones anteriores de JENUI, se clasifica como fundamental que el equipo docente adquiera las competencias necesarias para integrar la IA de manera efectiva en sus prácticas pedagógicas.

Este artículo presenta una experiencia de innovación cuyo objetivo es mejorar el aprendizaje del estudiantado mediante el uso de herramientas de IA generativa. En marzo de 2023, meses después del lanzamiento de ChatGPT basado en GPT-3.5, el equipo docente de la asignatura Sistemas Inteligentes II del grado de Informática en la Universidad de Málaga comenzó a realizar una actividad orientada a mejorar un resultado clave de aprendizaje: la capacidad de modelar problemas reales con redes bayesianas. Esta capacidad supone un cierto desafío para una buena parte del estudiantado, debido a su complejidad conceptual y técnica. Para ello, se les pidió que generaran sus propios enunciados utilizando IA generativa. Aunque existen estudios sobre el uso de estas herramientas para fomentar la generación de enunciados de problemas y así promover un aprendizaje personalizado [6], hasta donde conocemos las autoras y el autor de este artículo, no se han documentado experiencias donde sean los propios estudiantes los que generen dichos enunciados.

En septiembre de 2023, una de las profesoras del equipo docente de esta asignatura asistió al curso de innovación educativa impartido en la Universidad de Málaga titulado “Introducción al uso de la IA Generativa en la docencia universitaria”.

En una de las actividades de dicho curso, se propuso trabajar en equipo en el diseño de una actividad en la que el estudiantado utilizara las herramientas de IA generativa en su proceso de aprendizaje. Para el diseño se propuso utilizar una plantilla concreta, en la que se detallaban una serie de apartados. El equipo de trabajo se basó en la actividad desarrollada en la asignatura en el curso anterior para crear una ficha de actividad genérica, cuyo objetivo era abordar la resolución de tareas complejas con el apoyo de las herramientas de IA generativa. El objetivo era proporcionar un marco adaptable que permita al estudiantado comprender, estructurar y resolver problemas de alta dificultad con el uso de herramientas de IA generativa, fomentando un aprendizaje más autónomo y significativo.

Finalmente, durante el curso 23-24, se realizó de nuevo la actividad de uso del ChatGPT para aprender modelado con redes bayesianas, enriquecida con algunos nuevos elementos aportados por la concreción de los apartados de la ficha de actividad genérica desarrollada en el curso.

Para evaluar el impacto de la experiencia en los dos años de implantación, se realizó un análisis comparativo de las puntuaciones obtenidas en la pregunta de modelado de los exámenes en los cursos en los que se realizó la experiencia y en el inmediatamente anterior (tres cohortes, en total), observándose mejoras significativas en la adquisición de dicha competencia.

El objetivo de este artículo es compartir tanto el proceso de elaboración de la ficha de actividad genérica como los resultados de su aplicación en un caso concreto, con el fin de ofrecer una herramienta útil para la comunidad educativa y reflexionar sobre el potencial transformador de la IA generativa en el aprendizaje de tareas complejas, abundando en la necesidad de reflexión compartida en [7].

El artículo se estructura del siguiente modo: en el apartado II se describe el proceso de elaboración de una ficha de actividad genérica basada en IA generativa. En el apartado III-A se detalla el diseño de una actividad concreta a partir de dicha ficha, mientras que en el apartado III-B se presentan los resultados obtenidos. Finalmente, en los apartados IV y V se abordan otros resultados preliminares, así como la discusión general y las conclusiones del estudio.

II. ELABORACIÓN DE UNA FICHA DE ACTIVIDAD GENÉRICA

En este apartado presentaremos algunos aspectos relativos a la elaboración de la ficha de actividad genérica.

II-A. Contexto del curso de innovación educativa

El curso “Introducción al uso de la IA Generativa en la docencia universitaria”, impartido por Juan José Escribano Otero y María José García García en septiembre de 2023, constó de dos sesiones presenciales de dos horas, una sesión virtual síncrona de igual duración y cuatro horas de trabajo en grupo no presencial.

Su objetivo principal fue capacitar al profesorado en el aprovechamiento de las ventajas y posibilidades que ofrecen las he-

rramientas de IA generativa en el ámbito universitario, fomentándose la utilización dichas herramientas, entre otras cosas, para el apoyo al aprendizaje. La metodología fue activa y se propuso al profesorado diseñar, en grupo, una actividad evaluable en la que el estudiantado empleara IA generativa. Para ello, se proporcionó una plantilla común que aseguraba coherencia entre las actividades, enfocadas a resultados de aprendizaje, con tareas claras, aplicables a distintas disciplinas, generadoras de evidencias evaluables y acompañadas de técnicas de evaluación resistentes a la IA. Esta plantilla, basada en una propuesta previa para diseñar acciones formativas universitarias [8], promueve el diseño de actividades centradas en resultados de aprendizaje, contemplando recursos, métodos de evaluación y una reflexión previa sobre la viabilidad en términos de tiempo y ejecución, evitando así dificultades durante su implementación.

II-B. Ficha de actividad genérica elaborada en el curso

El equipo de trabajo formado en el marco del curso de innovación educativa estuvo integrado por seis docentes: una profesora de la ETSI de Telecomunicación y cinco docentes (tres profesoras y dos profesores) de la ETSI de Informática. Como se mencionó anteriormente, el diseño de la ficha de actividad genérica se basó en una experiencia previa desarrollada en la asignatura *Sistemas Inteligentes II* durante el curso anterior. Con el fin de organizar la información de forma clara, la Tabla 1 presenta: en la primera columna, los apartados y descriptores de la plantilla; en la segunda, los elementos de la ficha genérica elaborada por el equipo; y en la tercera, su aplicación concreta a una asignatura y actividad específica, la cual se detalla en la sección siguiente.

III. ELABORACIÓN Y PUESTA EN MARCHA DE UNA ACTIVIDAD CONCRETA: MODELADO CON REDES BAYESIANAS

Para ilustrar el uso de la ficha de actividad genérica, describiremos su aplicación a una tarea compleja concreta, el modelado con redes bayesianas.

III-A. Modelado con redes bayesianas

Las redes bayesianas [9] son modelos probabilísticos utilizados para modelar sistemas complejos y poder razonar con ellos en condiciones de incertidumbre. Se basan en la representación mediante un grafo acíclico dirigido de un conjunto de variables (nodos) y sus dependencias (enlaces). Entre sus ventajas se encuentran la capacidad de razonamiento que ofrecen, su bajo consumo en recursos computacionales, la facilidad de implementación, y, sobre todo, que son transparentes en cuanto a cómo procesan la información y generan los resultados, incluyéndose así dentro de los modelos de IA explicables, siendo intuitivas y comprensibles incluso para usuarios no técnicos. Su impacto en aplicaciones reales se extiende hasta la actualidad y le ha valido a su creador, Judea Pearl, numerosos premios, entre ellos el Premio Turing en el año 2011 y, más recientemente, el nombramiento como *Honorary Fellow* de la *Royal Society* en el año 2020.

TABLA I
FICHA DE ACTIVIDAD.

FICHA DE ACTIVIDAD		
APARTADOS y DESCRIPTORES	ACTIVIDAD PROPUESTA POR EL GRUPO EN EL CURSO	ACTIVIDAD IMPLANTADA EN LA ASIGNATURA
1.- Descripción detallada de la actividad del estudiantado con herramientas de IA generativa		
Denominación de la actividad: <i>Indica un nombre descriptivo para la actividad</i>	Evaluación, corrección y mejora de soluciones para tareas complejas con herramientas de IA generativa.	Modelado con redes bayesianas
Titulación y asignatura:	Cualquiera en la que se aborden tareas complejas	Sistemas Inteligentes II
Objetivos o resultados de aprendizaje <i>¿Qué se espera que el alumnado aprenda o sea capaz de hacer tras realizar esta actividad? Relacionado con los objetivos o resultados de aprendizaje de la asignatura/materia</i>	O1: Fomentar el autoaprendizaje de la materia específica.	O1: Fomentar el autoaprendizaje de la habilidad de modelar situaciones reales con redes bayesianas
Objetivos al introducir la IA en la actividad <i>¿por qué van a utilizar IA? ¿Qué se espera que el alumnado aprenda o sea capaz de hacer tras realizar esta actividad, específicamente relacionado con la utilización de la IA?</i>	O1: Fomentar el autoaprendizaje de la materia específica. O2: Fomentar el espíritu crítico O3: Comprobar la utilidad de las herramientas de IA generativa para ese tipo de tareas complejas	Respecto a los objetivos O2 y O3, en esta asignatura no se marcaron objetivos explícitos respecto al uso de la IA
Descripción de la actividad <i>Incluir la descripción de la actividad y su dinámica.</i>	Se seleccionará una tarea compleja relacionada con los resultados de aprendizaje de la asignatura y se propondrá al estudiantado que la resuelva con la ayuda de las herramientas de IA generativa	Se trata de utilizar la IA generativa para generar un enunciado sobre una situación de la vida real en la que pueda haber varias causas, efectos, indicios, etc. Posteriormente, el estudiantado deberá desarrollar un modelo de redes bayesianas para dicha situación
2.- Descripción del trabajo o tareas a realizar y estimación de tiempos		
Trabajo/tareas del profesorado <i>(Listado de tareas realizadas para preparar la actividad, durante la realización de esta y en la evaluación posterior)</i>	<ul style="list-style-type: none"> - Selección de la tarea compleja a resolver con la IA generativa - Generación del enunciado, incluyendo una propuesta de solución (que se mostrará o no al estudiantado) - Diseño de las tareas a entregar y de la rúbrica de evaluación de la tarea compleja. - Diseño de las encuestas pre-post tarea - Análisis de las encuestas pre-post tarea - Evaluación de las entregas del estudiantado. 	<ul style="list-style-type: none"> - Generación del enunciado y propuestas de solución - Diseño de las tareas a entregar - Configuración y administración de los foros de preguntas y respuestas - Corrección de las entregas y retroalimentación
Estimación del tiempo de trabajo para el profesorado <i>¿Cuánto tiempo se ha de emplear? Si se repite, ¿la siguiente vez costará más o menos?</i>	<ul style="list-style-type: none"> - Preparación de la tarea: 5 a 10 horas la primera vez, 1 a 2 horas las siguientes - Seguimiento del estudiantado: de 15 a 30 minutos por estudiante (corrección, retroalimentación, etc.). 	<ul style="list-style-type: none"> - Preparación de la tarea: 4 horas la primera vez, 2 las siguientes - Supervisión, evaluación de las entregas y generación de la retroalimentación: 15 minutos por estudiante
Trabajo/tareas del estudiantado dentro de clase <i>Listado de tareas dentro de las sesiones programadas de la asignatura, sólo o con sus compañeros si es un trabajo en equipo.</i>	<ul style="list-style-type: none"> - Pre-tarea: aprendizaje de las técnicas de resolución de la tarea compleja, familiarización con las herramientas de IA generativa - Post-tarea: resolución de tareas similares sin la ayuda de las herramientas de IA generativa 	El trabajo es individual y consiste en: <ul style="list-style-type: none"> - Pre-tarea: aprendizaje de las técnicas de resolución de la tarea compleja, familiarización con las herramientas de IA generativa - Post-tarea: resolución de tareas similares sin la ayuda de las herramientas de IA generativa
Trabajo/tareas del estudiantado fuera de clase <i>Listado de tareas fuera de las sesiones programadas de la asignatura, sólo o con sus compañeros si es un trabajo en equipo.</i>	<ul style="list-style-type: none"> - Generar la solución a la tarea compleja - Evaluarla, corregirla y mejorarla - Elaborar el entregable (basado en la rúbrica) - Responder a las encuestas pre-post 	<ul style="list-style-type: none"> - Generar el enunciado y la solución - Elaborar el entregable - Enviar la tarea por el foro de preguntas y respuestas

Estimación del tiempo de trabajo para el estudiantado <i>Indica tanto el tiempo empleado en la actividad durante las sesiones programadas de la asignatura como una estimación del tiempo adicional de trabajo independiente del estudiantado</i>	<ul style="list-style-type: none"> - Dentro de clase: 1 sesión (2 horas) para pre-tarea y 1 sesión (2 horas) para post-tarea - Fuera de clase: entre 4 y 6 horas (dependiendo de la complejidad de la tarea) 	<ul style="list-style-type: none"> - En clase: dos horas para aprender la habilidad de modelado - Fuera de clase: 2 horas
3.- Documentos y herramientas necesarias (Incluye herramientas de IA generativa y otras que se necesiten)		
Documentos y aplicaciones para el estudiantado	<ul style="list-style-type: none"> - Enunciado de la tarea - Rúbrica - Herramientas de IA generativa apropiadas al tipo de tarea compleja que se presente - Documentación propia de la materia / asignatura - Software de prácticas específico de la materia/asignatura 	<ul style="list-style-type: none"> - Apuntes de la asignatura - Enunciado de la tarea - Herramientas de IA generativa a elección del estudiantado - Foros del Campus Virtual - GeNIe (herramienta software de libre uso para redes bayesianas)¹[10]
Documentos y aplicaciones para el profesorado	<ul style="list-style-type: none"> - Materiales del curso ““Introducción al uso de la IA Generativa en la docencia universitaria”” - Herramientas de IA generativa 	<ul style="list-style-type: none"> - GeNIe (software específico de la materia) - Foros del Campus Virtual
Condiciones necesarias para desarrollar la actividad	<ul style="list-style-type: none"> - Laboratorios docentes - Software específico de la materia - Campus Virtual 	<ul style="list-style-type: none"> - GeNIe - Campus Virtual
4.- Evaluación:		
Resultados de aprendizaje: <i>¿Qué vas a evaluar y cómo?</i>	<ul style="list-style-type: none"> - Resultados de aprendizaje pertinentes de la asignatura - Capacidad para resolver la tarea compleja 	<p>RA3. Representar el conocimiento de un modo computable</p> <p>RA4. Conocer y aplicar las técnicas de razonamiento bajo incertidumbre y bajo imprecisión</p> <p>Capacidad para modelar situaciones de la vida real con redes bayesianas</p>
Instrumentos de evaluación: <i>¿Qué tipo de instrumentos vas a utilizar? (Examen, Exposición oral, Rúbrica, Informe escrito)</i>	<p>E1. Rúbrica de la tarea</p> <p>E2. Examen</p> <p>E3. Encuestas pre-post tarea</p> <p>E4. Informe reflexivo sobre las conclusiones de la tarea</p>	E2. Examen
Indicadores y tipo de evaluación (formativa/sumativa) <i>¿Qué tipo de evaluación vas a aplicar para confirmar la consecución de los objetivos o resultados de aprendizaje definidos para la actividad en el apartado 1 de esta ficha? (Autoevaluación / coevaluación / evaluación formativa /sumativa...)</i>	<p>Para O1:</p> <ul style="list-style-type: none"> - Indicadores específicos de la materia para ser incluidos en la rúbrica (formativa) y ser evaluados en el examen (sumativa) <p>Para O2 y O3</p> <ul style="list-style-type: none"> - Validez del análisis realizado sobre adecuación de los productos generados por las herramientas de IA generativa para la resolución de la tarea - Profundidad y adecuación de la reflexión del estudiantado 	<p>Para O1:</p> <p>Los indicadores que se miden son:</p> <p>Adecuación de la definición de los nodos y su correspondiente estado; de las relaciones de influencia causal y de los modelos de interacción entre los nodos (relación de AND, OR, etc.)</p> <p>En la tarea se miden para realizar evaluación formativa, en el examen para realizar evaluación sumativa</p>
Evaluación en convocatoria extraordinaria y otras situaciones	E4. Examen (Esta actividad no se realizará en evaluación extraordinaria)	E4. Examen (Esta actividad no se realizará en evaluación extraordinaria)

Llamamos *modelado con redes bayesianas* al proceso de ser capaz de elaborar un modelo que describa correctamente una situación, problema o sistema real. Para realizar el modelado de la red, se debe en primer lugar definir las variables de esta, que tienen que tomar un conjunto exhaustivo y excluyente de valores. Una vez definidas las variables, hay que determinar las relaciones de influencia causal entre ellas, de forma que el

grafo final resultante sea acíclico y dirigido. Y, por último, se han de estimar los parámetros del modelo, en este caso, las tablas de distribución de las probabilidades condicionadas de cada variable dados sus padres. El proceso de modelado es el más importante a la hora de usar las redes bayesianas como formalismo de razonamiento para un sistema complejo, ya que

¹ <https://www.bayesfusion.com/genie/> GeNIe Modeler. BayesFusion, LLC..

una vez definida la red, existen multitud de aplicaciones y librerías para razonar con ellas.

Aunque este proceso pueda parecer sencillo, nuestra experiencia es que resulta ciertamente complejo para buena parte del estudiantado de la asignatura, que normalmente tienen problemas en las dos primeras etapas del proceso (definición de variables e identificación de relaciones de influencia causal). Más concretamente, a la hora de definir las variables no tienen en cuenta que deben formularse de forma que sus valores sean exhaustivos y excluyentes, y a la hora de definir los enlaces, tienen dificultades serias para interpretar las relaciones causales. Ilustremos dichas dificultades con un ejercicio que usaremos a modo de ejemplo (ver Figura 1).

Ejercicio: Considera la siguiente situación: en el planeta ZYX se pueden encontrar varias clases de animales, llamemos a estas clases Wurros, Hobexas y Wackas. Todos tienen un tamaño muy pequeño, y sus pieles son o bien escamosas o bien están cubiertas de suave pelo. Además, una observación atenta ha permitido deducir lo siguiente. Todos los Wurros tienen 5 ó 6 patas. Su color es rojizo, y tienen la piel peluda y suave. El número de patas de las Hobexas es un entero que varía uniformemente entre 4 y 6, ambos inclusive. Su piel es escamosa. En cuanto a las Wackas, tienen 4 o 5 patas, y ofrecen a la vista una tonalidad casi siempre azulada, pero a veces (20% de los casos) rojiza. Los animales que tienen un número impar de patas cojean siempre. Los animales que tienen un número par de patas cojean sólo cuando tienen alguna anomalía (malformación congénita, heridas, etc.), lo cual ocurre en el 10% de los casos para los animales de 4 patas, y en el 20% para los de seis. Plantea el problema de clasificación de los animales de ZYX mediante una red bayesiana.

Figura 1. Ejemplo de problema de modelado

Para entender la complejidad del problema, quizás el lector o lectora quiera intentar elaborar su propio modelo de redes bayesianas en este momento. Para resolver correctamente el problema, y, teniendo en cuenta que las variables deben definirse de forma que el conjunto de sus estados sea exhaustivo y excluyente, se podrían definir por ejemplo las siguientes variables y estados: *Tipo de animal*, con valores *wurro*, *wacka* y *hobexa*; *Número de patas*, con valores 4, 5 y 6; *Color*, con valores *rojizo* y *azulado*; y *Cojea*, con valores *sí* y *no*. Para los enlaces, se ha de tener en cuenta que es el *tipo de animal* el que tiene influencia causal en las diferentes características, y que, según el enunciado, *el número de patas* tiene influencia causal en que el animal cojee o no. En la Figura 2 se muestra el grafo asociado a esta solución correcta a este ejercicio

Por supuesto, el mismo problema puede resolverse correctamente de otras maneras. Por ejemplo, en relación con la anomalía, la solución mostrada en la Figura 2 la contempla al definir las probabilidades condicionadas del nodo *Cojea*. Sin embargo, también sería válido modelarla mediante una variable específica denominada *Anomalía*, que actuaría como nodo intermedio en el grafo: hijo del nodo *Número de patas* y padre del nodo *Cojea*.

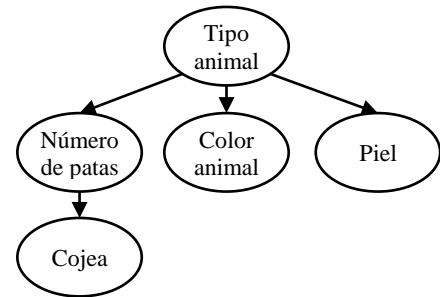


Figura 2. Una solución correcta para el problema de modelado

Pese a que en la clase de modelado se insiste mucho en estos aspectos, planteando ejemplos concretos para que vayan pensando y resolviendo, una cantidad significativa de estudiantes comete errores de concepto en este ejercicio (que es el primero que resuelven de manera autónoma). De este modo, un número significativo propone soluciones parecidas a la que se muestra en la Figura 3 (en la que, por simplicidad, solo mostramos los nodos relativos al *tipo de animal* y las características *número de patas* y *color*):

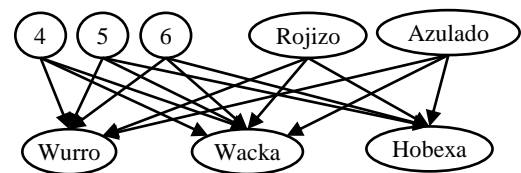


Figura 3. Una solución incorrecta para el ejercicio

Como se puede ver, la solución incorrecta tiene dos tipos de errores. En primer lugar, definir como variables (nodos) lo que deben ser valores de las variables. En segundo lugar, interpretar incorrectamente la relación de causalidad entre las variables, debido probablemente a que se está pensando en reglas de diagnóstico (por ejemplo, si el animal tiene 4 patas es “porque” es un *wurro*) más que en relaciones de influencia causal (el tipo de animal influye en sus características). Esta dificultad para abordar problemas de modelado se ve reflejada también en la calificación que obtienen en el problema que se incluye en las pruebas escritas, que suele ser bastante inferior a la obtenida en otras preguntas que son de realización más mecánica.

Por todo ello, en los dos últimos cursos académicos se ha desarrollado una actividad específica para que el estudiantado pueda mejorar sus habilidades de modelado, que pasamos a describir a continuación.

III-B. Actividad implantada en la asignatura

La asignatura se imparte en tercer curso del grado de Informática en la Universidad de Málaga. Cuenta con más de cien estudiantes, y la docencia se imparte en dos grupos (A y B). El enunciado completo de la actividad propuesta en el curso 22-23 fue la que se muestra en la Figura 4.

Tal como ya se ha indicado, esta actividad en la que tenían que usar alguna herramienta de IA generativa se propuso por primera vez en el curso 22-23, con el objetivo de mejorar las destrezas de modelado del estudiantado, lo cual podríamos

comprobar a través de las calificaciones en el problema de modelado que se incluye en el examen de la asignatura. La Figura 5 muestra una solución correcta propuesta por un estudiante en el curso 22-23.

En esta tarea se pide elaborar un enunciado de un problema de modelado de una situación real (simplificada) con una red bayesiana. Para ello sigue los siguientes pasos:

1. Identifica un dominio o problema de tu interés. Para ello, busca información con alguna herramienta de IA generativa sobre algún tema concreto.
2. Marca en el texto original la información que consideres relevante para ese dominio: síntomas, indicios, pruebas, causas (potenciadoras, inhibidoras, favorecedoras, preventivas, etc.).
3. Simplifica el problema para eliminar posibles ambigüedades de las relaciones de influencia causal y tener un número de nodos razonable para un problema de clase (entre 8 y 12)
4. Escribe el enunciado.
5. Crea el grafo correspondiente a tu red (no olvides etiquetar las relaciones AND/OR, procura que la red tenga al menos un tipo de relación de cada una de ellas).
6. Sube al foro del campus virtual tu solución, incluyendo los siguientes elementos: título, referencia original, texto original, enunciado, imagen del grafo de la red bayesiana.

Figura 4. Enunciado de la actividad propuesta en el curso 22/23

Título: Derrota del SevillaFC

Referencia: Chat GPT

Yo: Cuales son las causas de que el Sevilla FC tenga una derrota

ChatGPT: Las causas de una derrota del Sevilla FC pueden variar y pueden ser el resultado de una combinación de factores. Algunas posibles causas incluyen:

1. Rendimiento del equipo: El bajo rendimiento de los jugadores en el campo, ya sea debido a la falta de concentración, errores individuales, falta de coordinación o problemas tácticos, puede llevar a una derrota.
2. Estrategia del oponente: El rival puede haber desarrollado una estrategia efectiva para contrarrestar el estilo de juego del Sevilla FC, aprovechando sus debilidades y capitalizando en ellas.
3. Lesiones o suspensiones de jugadores clave: La ausencia de jugadores importantes debido a lesiones, suspensiones u otras razones puede afectar negativamente el rendimiento del equipo y aumentar las posibilidades de una derrota.
4. Fatiga física o mental: El cansancio acumulado debido a una temporada larga, viajes frecuentes, partidos consecutivos u otros factores puede afectar el rendimiento de los jugadores y contribuir a una derrota.
5. Decisiones arbitrales controvertidas: Decisiones arbitrales polémicas pueden influir en el resultado del partido y llevar a la derrota del Sevilla FC.
6. Factores externos: Condiciones climáticas adversas, estado del terreno de juego o apoyo de la afición local pueden ser factores externos que contribuyan a una derrota.

Enunciado: Los aficionados del SevillaFC siguen preguntándose cuáles son las posibles causas de que su equipo siga perdiendo partidos. Los expertos dicen que puede ser causa de la fatiga de la plantilla por el partido reciente contra el Atlético de Madrid y por el poco descanso que han tenido los jugadores entre partidos. También aseguran que la suspensión de En-Nesyri, delantero clave del club, ha provocado una alineación inadecuada de cara a los encuentros. La afición también es consciente del descontento dentro del equipo por problemas de la directiva y las constantes derrotas anteriores.

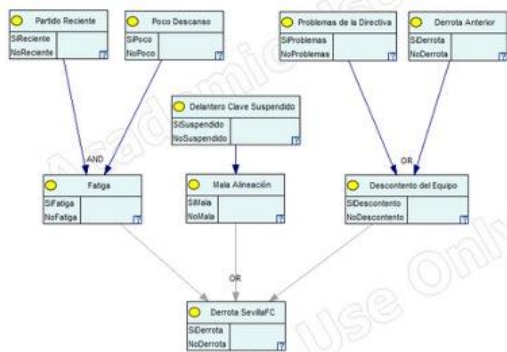


Figura 5. Solución correcta propuesta por un estudiante

Tras la participación en el curso y elaboración de la ficha de actividad genérica, en el curso 23-24 se elaboró la ficha de actividad de modelado con redes bayesianas mostrada en la tercera columna de la Tabla 1, que se llevó a cabo en ese mismo curso, con algunas novedades. En primer lugar, en el grupo A se dio retroalimentación sobre la tarea (evaluación formativa), calificando cada modelo como “mejorable” (con una indicación de los fallos, tales como definición incorrecta de variables o direcciones de los enlaces, etc.) o “correcto”. A partir de estos comentarios corrigieron sus modelos y generaron una extensa relación de problemas resueltos. En segundo lugar, como actividad adicional, cada estudiante tuvo que utilizar su enunciado para trabajar sobre él todos los conceptos teóricos relativos a redes bayesianas (especificación de parámetros, propiedades independencia condicional, teorema de factorización, evolución de las probabilidades en la red conforme se adquieren nuevas evidencias y algoritmos de propagación). En el grupo B de la asignatura no se dio retroalimentación sobre el problema de modelado, para poder evaluar la influencia de esta en los resultados finales. De esta forma, la nota de la pregunta en el examen permitirá también evaluar si el esfuerzo realizado para corregir las tareas entregadas por el estudiantado (evaluación formativa) tiene impacto en el aprendizaje (demostrado en los resultados de la evaluación sumativa

La Figura 6 muestra el grafo de una solución incorrecta dada por un estudiante del grupo A en el curso 23-24, y la correspondiente retroalimentación proporcionada por el equipo docente de la asignatura.

```

    graph TD
      T[Tos] --> S[Síntoma]
      F[Fiebr] --> S
      L[Llevar mascarilla] --> P[Prevenciones]
      D[Distancia seguridad] --> P
      S --> COVID[COVID]
      P --> COVID
      COVID --> FA[Falta aire]
      COVID --> PO[Pérdida olfato]
  
```

Promedio de valoraciones: mejorable (1) mejorable 2

No se ha enviado por email Enlace permanente Mostrar intervención inicial Divide Borrar Responder

y algunos ejemplos de respuesta

miércoles, 6 de marzo de 2024, 17:50

Hola, en esta red hay cosas importantes que corrigir: fíjate que tienes cuatro síntomas (tos, fiebre, pérdida de olfato, falta de aire) y dos están correctamente como hijos de la enfermedad que los produce (COVID) mientras que otros los has puesto como causas de la enfermedad. Deben estar también como hijos de COVID. Respecto a llevar mascarilla y mantener la distancia, son factores preventivos y está bien que el quien los agarre en un nodo que sea padre de COVID, ya que tomar esas precauciones disminuirá la probabilidad de contraerlo. Faltaba entonces poner alguna causa para COVID, como pueden ser tener contacto con el virus (sería padre del nodo) y entonces te podría quedar una estructura tipo AND arriba (para ello tendrías que poner el nodo prevenciones negado, es decir, tener contacto con el virus AND no tomar prevenciones provoca que cojas el COVID). También podrías meter Resfriado como otra posible enfermedad, ya que la mencionas en el texto. Te recomiendo que entres la red con estos comentarios y respondas con la nueva imagen a este comentario.

Promedio de valoraciones: -

No se ha enviado por email Enlace permanente Mostrar intervención inicial Divide Borrar Responder

Figura 6. Solución incorrecta propuesta por un estudiante y retroalimentación proporcionada por el profesorado

Por tanto, la experiencia realizada genera tres situaciones diferentes, en base a la cual podemos categorizar al estudiantado en tres grupos: C1-sin actividad (106 estudiantes del curso 21-22); C2-con actividad, pero sin retroalimentación (114 estudiantes del curso 22-23 y 64 estudiantes del grupo B del curso 23-24); C3 con actividad y con retroalimentación (62 estudiantes del grupo A del curso 23-24).

En base a dicha categorización se han generado los grupos para el estudio: el grupo G1, correspondiente a la cohorte 21-22, que no realizó la tarea de modelado; el grupo G2, de la cohorte 22-23, que incorporó dicha tarea sin intervenciones

adicionales; y dos grupos del curso 23-24, G3 y G4, ambos realizando la tarea de modelado, pero con la diferencia de que G3 recibió retroalimentación directa del profesorado, mientras que G4 no contó con este apoyo adicional. Para una mejor visualización de estos grupos y su distribución temporal, se puede consultar las dos primeras columnas de la tabla II.

TABLA II
RESULTADOS PARA LAS TRES CONDICIONES.

Curso	Número estudiantes y calificación media		
	Condición	Nº estudiantes	Calificación
21-22	G1-sin actividad	106	6,85
22-23	G2-con actividad	114	9,65
23-24	G3-con actividad	64	8,99
	G4-con actividad y realimentación	62	9,07

Las comparaciones se llevaron a cabo en dos etapas: en primer lugar se compararon los grupos G1 y G2 para determinar si la implementación de la tarea tuvo un impacto significativo en las calificaciones y, en segundo lugar, se realizaron las comparaciones entre los grupos G3 y G4 para analizar el efecto de la retroalimentación en el rendimiento académico.

Como ya se ha comentado, para evaluar el impacto de la actividad, se usó la calificación en la pregunta de modelado de las pruebas finales escritas (examen). En el apartado IV presentaremos y analizaremos los resultados.

IV. ANÁLISIS DE RESULTADOS

IV-A. Resultados e interpretación

La Tabla II muestra el número de estudiantes que participó en cada condición, junto con la calificación media obtenida en la pregunta de modelado del examen (indicador I2), escalada a 10 puntos.

Para visualizar las diferencias entre los grupos, se realizó un análisis preliminar mediante un diagrama de cajas y bigotes, (*boxplot*) que se muestra en la Figura 7. En el eje horizontal se indican cada uno de los grupos y en el vertical, las calificaciones alcanzadas, en escala de 0 a 10.

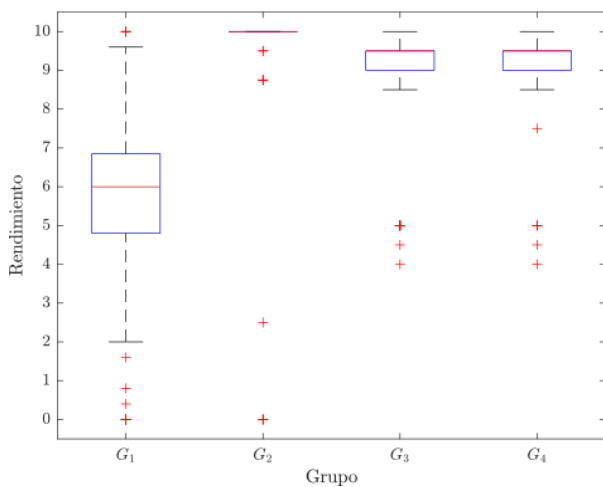


Figura 7. Diagrama de cajas y bigotes de las calificaciones de los grupos.

Podemos observar que los grupos que realizaron la tarea de modelado (G2, G3 y G4) presentan una mejor media y menor dispersión, que indica un rendimiento académico superior en comparación con el grupo G1, que no realizó la tarea.

Para visualizar las diferencias entre los grupos, se realizó un análisis preliminar mediante un diagrama de cajas y bigotes, (*boxplot*) que se muestra en la Figura 7. En el eje horizontal se indican cada uno de los grupos y en el vertical, las calificaciones alcanzadas, en escala de 0 a 10.

Podemos observar que los grupos que realizaron la tarea de modelado (G2, G3 y G4) presentan una mejor media y menor dispersión, que indica un rendimiento académico superior en comparación con el grupo G1, que no realizó la tarea.

IV-B. Análisis de los datos

El análisis de los datos obtenidos se centró en evaluar el impacto de la tarea de modelado con redes bayesianas y el efecto de la retroalimentación proporcionado por los profesores de la asignatura sobre el rendimiento académico del estudiantado. Se optó por utilizar la prueba no paramétrica de Mann-Whitney U, una herramienta estadística robusta que permite comparar distribuciones de dos grupos independientes sin requerir la asunción de normalidad en los datos, lo que asegura la validez de los resultados en contextos educativos donde las distribuciones pueden ser asimétricas o contener valores atípicos.

En la comparación entre G1 y G2, los resultados indicaron una mejora sustancial en las calificaciones tras la introducción de la tarea. La media de las calificaciones de G1 fue de 6.85 (± 2.71), mientras que la de G2 alcanzó 9.60 (± 1.81).

El análisis mediante la prueba de Mann-Whitney U [10] arrojó un valor p de 1.73×10^{-21} , lo que permite rechazar la hipótesis nula y confirmar que las diferencias en las distribuciones son estadísticamente significativas. Además, los *rankings* medios refuerzan este hallazgo, siendo de 67.81 para G1 y 143.83 para G2.

Este resultado evidencia que la introducción de la tarea de modelado tiene un efecto positivo claro y significativo en el aprendizaje que el estudiantado demuestra en el examen, lo que sugiere que esta actividad basada en el uso de IA generativa favorece la comprensión de conceptos complejos, como el modelado con redes bayesianas.

En la segunda comparación, entre los grupos G3 y G4, se buscó evaluar el impacto de la retroalimentación proporcionado por los profesores de la asignatura. Las medias de las calificaciones fueron de 8.99 (± 1.32) para G3 y 9.01 (± 1.71) para G4, lo que indica que ambos grupos alcanzaron resultados similares. Por otra parte, el *ranking* medio de G3 fue ligeramente superior al de G4 (71.23 frente a 70.73), lo que sugiere una mejor posición relativa de las personas del grupo al que se le dio retroalimentación. La mayor media y el mejor *ranking* medio de G3 podrían interpretarse como una leve indicación de que la retroalimentación refuerza la comprensión del estudiantado, permitiéndole abordar con mayor profundidad los problemas planteados, si bien la prueba de Mann-Whitney U no detectó diferencias estadísticamente significativas en esta comparación (valor p = 0.937).

En resumen, los resultados del análisis evidencian que integrar tareas de modelado con apoyo de IA generativa tiene un efecto positivo en el aprendizaje. Esta metodología no solo mejora el rendimiento académico, sino que facilita la comprensión de contenidos complejos, como el modelado con redes bayesianas. Aunque no se observaron diferencias estadísticamente significativas entre los grupos G3 y G4, la tendencia apunta a que la retroalimentación docente puede contribuir a consolidar el aprendizaje, lo que subraya el valor de combinar tecnologías avanzadas con la interacción educativa directa.

IV-C. Otros resultados preliminares

Durante el curso 23-24, se implementó además una nueva actividad en la que cada estudiante eligió un problema personalizado como base para todas las actividades teóricas sobre redes bayesianas, en lugar de emplear problemas estandarizados. A partir de estos casos individuales, aplicaron conceptos clave como la independencia condicional, el teorema de factorización, algoritmos básicos de propagación de probabilidades y el análisis de la actualización de creencias ante nueva evidencia.

Estas competencias, incluidas entre los resultados de aprendizaje evaluables de la asignatura, han mostrado una mejora preliminar durante el curso 23-24, aunque aún se requiere una evaluación más exhaustiva. También se han identificado beneficios subjetivos relevantes. El equipo docente ha notado un aumento en la motivación e interés del alumnado por las redes bayesianas, al trabajar con problemas cercanos a su realidad y alineados con sus propios intereses (como deportes, inversiones, literatura, historia, jardinería o juegos). Esto ha favorecido un aprendizaje más personalizado, con soluciones más creativas y significativas. Además, el profesorado ha experimentado una mayor motivación, al conocer mejor a su estudiantado y sentirse más conectado con ellos. También ha percibido una mejor gestión del tiempo docente, al poder centrarse en tareas de mayor valor pedagógico, como ofrecer retroalimentación, en lugar de generar problemas individualizados para cada estudiante, una tarea que sería poco viable en grupos numerosos.

V. DISCUSIÓN Y CONCLUSIONES

Este artículo ha presentado una plantilla orientada al diseño de actividades educativas que integran herramientas de IA generativa con el objetivo de mejorar el aprendizaje (columna 1 de la Tabla I). Su aplicación facilita una mayor coherencia interna en la asignatura y permite adaptar el tiempo de dedicación del alumnado según los créditos ECTS asignados, en línea con lo señalado en [11], que destaca la necesidad de pautas claras para la integración eficaz de estas tecnologías en la educación.

A partir de esta plantilla, se propuso una ficha de actividad genérica para el desarrollo de tareas complejas (columna 2) y se expuso el diseño y la evaluación de una actividad concreta (columna 3). Esta ficha promueve una reflexión crítica sobre la elección de herramientas de IA generativa más adecuadas para cada actividad, y puede representar un paso riguroso y estructurado hacia la adaptación del proceso enseñanza-aprendizaje universitario, incluida la evaluación. Esta visión coincide con

investigaciones recientes como [12], que respaldan una integración cuidadosa de la IA generativa, siempre que se consideren sus riesgos, como el plagio o la dependencia tecnológica.

Respecto a su aplicabilidad en otras asignaturas, el modelo es replicable: otros equipos docentes pueden utilizar la Tabla I para diseñar sus propias fichas (columna 2), con objetivos distintos, o construir actividades específicas (columna 3) adaptadas a las características y necesidades de sus contextos formativos.

Los resultados obtenidos en la asignatura muestran una mejora clara en el aprendizaje del modelado con redes bayesianas, reflejada en las calificaciones más altas obtenidas por las cohortes que realizaron la actividad propuesta. Estos datos coinciden con estudios previos como [13], que evidencian mejoras académicas tras la incorporación de IA generativa en entornos educativos. No obstante, no se ha podido establecer con certeza que la retroalimentación docente haya influido directamente en las notas finales. Aunque no se han analizado formalmente las causas del escaso impacto observado de la retroalimentación, este efecto podría estar relacionado con otros factores. En particular, al consultar a otros equipos docentes de asignaturas del mismo curso, se observó que la nota media en sus asignaturas era ligeramente superior en el grupo B respecto al grupo A. Esto podría sugerir que el estudiantado del grupo B presentaba una mayor capacidad de aprendizaje, lo que le habría permitido alcanzar resultados similares sin haber recibido retroalimentación ni contar con la relación de problemas resueltos.

Entre las limitaciones del estudio, aunque el tamaño muestral (entre 62 y 114 estudiantes) no es reducido, sería recomendable ampliarlo para mejorar la validez y generalización de los resultados. Además, no se han controlado otros factores potencialmente influyentes, como la experiencia previa del alumnado con herramientas de IA, sus habilidades específicas en la materia, o su capacidad general de aprendizaje, o posibles diferencias en la dificultad de la pregunta de examen. Considerar estas variables en investigaciones futuras permitiría obtener una visión más completa y precisa, tal como se sugiere en [14].

En cuanto a líneas futuras, el equipo docente planea profundizar en el uso de problemas personalizados y evaluar su impacto de forma más sistemática, incluyendo encuestas al estudiantado. Otra posible línea de trabajo interesante sería realizar un estudio sistemático del diseño, uso e interpretación de los *prompts* utilizados por los estudiantes.

AGRADECIMIENTOS

Esta investigación ha sido financiada por el Ministerio español de Ciencia e Innovación a través del proyecto de investigación con referencia TED2021-129956B-I00.

REFERENCIAS

- [1] W. Holmes, M. Bialik y C. Fadel, "Artificial intelligence in education". En *Data ethics : building trust : how digital technologies can serve humanity*, Ginebra: Globethics Publications, pp 621-653, 2023. [doi: 10.58863/20.500.12424/4276068](https://doi.org/10.58863/20.500.12424/4276068).
- [2] R. Luckin y W. Holmes, *Intelligence unleashed: An argument for AI in Education*. Pearson, 2016.

- [3] O. Zawacki-Richter, V.I. Marín, M. Bond y F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education—where are the educators?" En *International Journal of Educational Technology in Higher Education*, 16, 39, 2019 doi: doi.org/10.1186/s41239-019-0171-0
- [4] M.A. Serrano, A. Santos-Olmo, D.G. Rosado, L. E. Sánchez, C. Blanco y E. Fernández-Medina. "Uso de ChatGPT en la enseñanza universitaria de informática: ventajas y desventajas" en *Actas de las JENUI*, pp. 233-240, 2024.
- [5] F.P. Romero, J. Serrano-Guerrero, J.A. López-Gómez, L. Jiménez y J. A. Martín-Baos. "Experiencia docente preliminar con chatGPT: desafíos y adaptaciones" en *Actas de las JENUI*, pp. 205-208, 2023.
- [6] M. Mudassar Yamin, E. Hashmi, M. Ullah y B. Katt, "Applications of LLMs for Generating Cyber Security Exercise Scenarios," in *IEEE Access*, vol. 12, pp. 143806-143822, 2024, doi: [10.1109/ACCESS.24.3468914](https://doi.org/10.1109/ACCESS.24.3468914).
- [7] L. Jiménez, J.A. López-Gómez, J.A. Martín-Baos, F.P. Romero y J. Serrano-Guerrero. "ChatGPT: reflexiones sobre la irrupción de la inteligencia artificial generativa en la docencia universitaria," en *Actas de las JENUI*, pp. 113-120, 2023.
- [8] M. J. García García, M. J. Terrón López y Y. Blanco Archilla. "Desarrollo de recursos docentes para la evaluación de competencias genéricas". En *Carlos Vaz de Carvalho, Ricardo Silveira, Manuel Caeiro Rodriguez, editores. TICAI2009: TICs para a Aprendizagem da Engenharia*. IEEE, Sociedad de Educación: Capítulos Español y Portugués. ISBN 978-972-8688-70-7, pp.139-144, 2009.
- [9] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, 1988.
- [10] H. B. Mann y D. R. Whitney. "On a test of whether one of two random variables is stochastically larger than the other" En *Annals of Mathematical Statistics*, 18, pp. 50–60, 1947.
- [11] N. McDonald, A. Johri, A. Ali y A. Hingle, "Generative Artificial Intelligence in Higher Education: Evidence from an Analysis of Institutional Policies and Guidelines" En *Computers in Human Behavior: Artificial Humans*, vol 3, 2025. <https://doi.org/10.1016/j.chbah.2025.100121>
- [12] J. E. Pérezchica-Vega, J. A. Sepúlveda-Rodríguez y J. P. Hernández-Ramos,. "Inteligencia artificial generativa en la educación superior: usos y opiniones de los profesores" En *European Public & Social Innovation Review*, 9, pp. 1–19, 2024. <https://doi.org/10.31637/epsir-24-593>
- [13] N. Yusuf, N. Alsulami y A. Ouda, "Generative AI and the Future of Higher Education: A Threat to Academic Integrity or Reformation? Evidence from Multicultural Perspectives". En *International Journal of Educational Technology in Higher Education*, 21(1), 10, 2024. <https://doi.org/10.1186/s41239-024-00453-6>
- [14] A. Briñis Zambrano. Beneficios y limitaciones en docentes y estudiantes universitarios salvadoreños sobre el uso de IA en procesos de enseñanza-aprendizaje. En *European Public & Social Innovation Review*, 9, pp. 1–19. 2024 <https://doi.org/10.31637/epsir-24-368>